

Title: *Geographic and Pricing Analysis of Indian Agricultural Market Data*

Objective:

To perform exploratory data analysis and segmentation on Indian agricultural market data to extract valuable insights related to **geographic regions** and **pricing behavior**. The analysis aims to benefit key stakeholders like **farmers**, **agri-businesses**, and **policy makers** by identifying:

- Regions (states/districts/markets) with high price volatility.
- Commodities with the highest average modal prices.
- Patterns in commodity availability and pricing across different geographies.
- Insights to improve decision-making in crop selection, market targeting, and supply chain optimization.

Dataset Features:

The dataset contains 9 columns and 2,238+ rows with the following attributes:

- `state`: Indian state name
- `district`: District within the state
- `market`: Specific market within the district
- `commodity`: Name of the agricultural commodity
- `variety`: Variety of the commodity
- `arrival_date`: Date of the commodity arrival at market
- `min_price`, `max_price`, `modal_price`: Recorded prices for the commodity on the arrival date

Goals:

1. Segment data based on geographic variables (`state`, `district`, `market`).

2. Analyze pricing patterns (min_price, max_price, modal_price).
3. Identify high-value crops and regions with significant pricing opportunities.
4. Visualize price distribution, volatility, and commodity presence across the country.

Analytical Steps

1. Data Preparation

The first step involves cleaning and preparing the dataset:

- Convert date formats for future time-based analysis.
- Create a new feature called **price spread**, which is the difference between the maximum and minimum price for each commodity entry. This helps identify volatility in pricing.

2. State-wise Commodity Distribution

We analyze how many unique commodities are present in each state.

This reveals the agricultural diversity of each region.

States with more diverse commodities might support multi-crop farming.

It indicates potential for local food systems and export.

Helps identify states that may benefit from agri-tech interventions.

Such information is key for regional development policies.

This also supports targeting for subsidies and training.

High diversity states could attract agribusiness investments.

3. Market Activity Analysis

This step identifies the most active markets in the dataset.

Active markets are those with high transaction frequency.

They serve as trade hubs for specific districts or states.

These markets require adequate infrastructure and policy focus.

Frequent markets may offer better price discovery for farmers.

They are also useful for setting up warehouses and cold chains.

Understanding market activity helps optimize supply chains.
Such analysis informs logistics, distribution, and investment.

4. Identifying High-Value Commodities

This analysis finds commodities with the highest average modal prices.
Such crops are likely to provide higher profits to farmers.
Understanding which commodities are most lucrative guides crop selection.
High-value crops can attract private investment and R&D.
It helps the government promote premium exports or support price stability.
Farmers can align with market demand and increase income.
This insight is useful for regional specialization strategies.
Such crops may also require risk management and insurance.

5. Top Performing States

This section identifies states with the highest average modal prices.
These states may benefit from better farming practices or infrastructure.
They might be closer to consumer markets or export ports.
High prices can reflect better supply chain efficiency.
It reveals where farmers earn more for similar produce.
Can guide future agricultural investments or public schemes.
This also highlights regional inequalities in pricing.
Policy adjustments may be required for underperforming states.

6. Price Volatility Analysis

Price volatility is measured using the spread between maximum and minimum prices.
States with high average spreads may face unstable market conditions, which could be risky for farmers and require policy intervention or better market regulation.

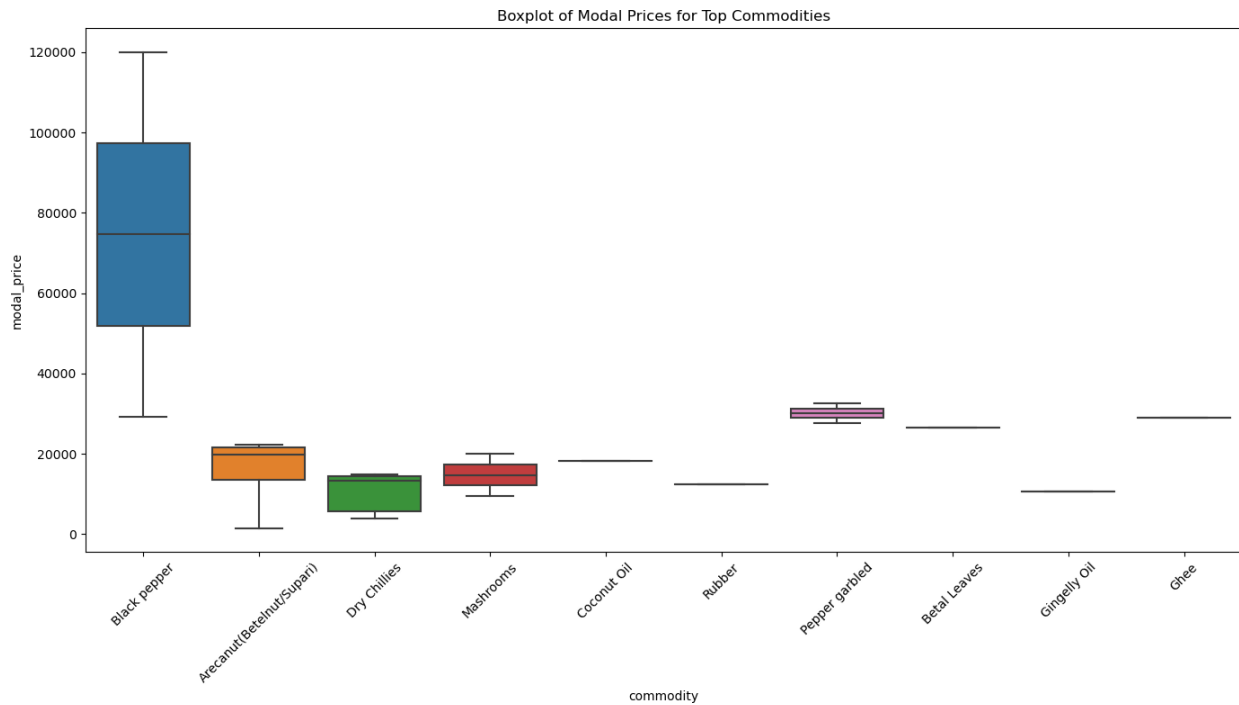
7. Heatmap Visualization: State vs. Commodity Prices

Policy adjustments may be required for underperforming states.



8. Price Distribution using Boxplots

Boxplots are used to show the distribution of prices for the most valuable commodities. This helps understand the consistency and variability of market prices, revealing commodities that might have unstable or highly fluctuating pricing patterns.



9. Choropleth Map: Pricing Across States

An interactive choropleth map displays the average commodity prices across Indian states. This helps in visually identifying regions with higher or lower agricultural market prices, useful for policy decisions and regional planning.

Insights & Applications

- **Farmers:** Identify high-return crops and the best markets for selling their produce.
- **Policymakers:** Detect regions with pricing instability or low market performance.
- **Startups & Agri-businesses:** Spot regional opportunities for investment, logistics, or value-chain improvements.

Recommendations for Future Work

- Incorporate **time-series trends** to track seasonal price changes.
- Use **clustering techniques** to create price-region segments for better targeting.
- Integrate **weather, soil, and rainfall** data for deeper agricultural forecasting.

Github

link:https://github.com/Manekuntaramesh/Market_Segmentation_Analysis/blob/main/Geographic%20and%20Pricing%20Analysis%20of%20Indian%20Agricultural%20Market%20Data%20.ipynb